



WP3 – Information Platform Progress Report

4th Epiwork Meeting

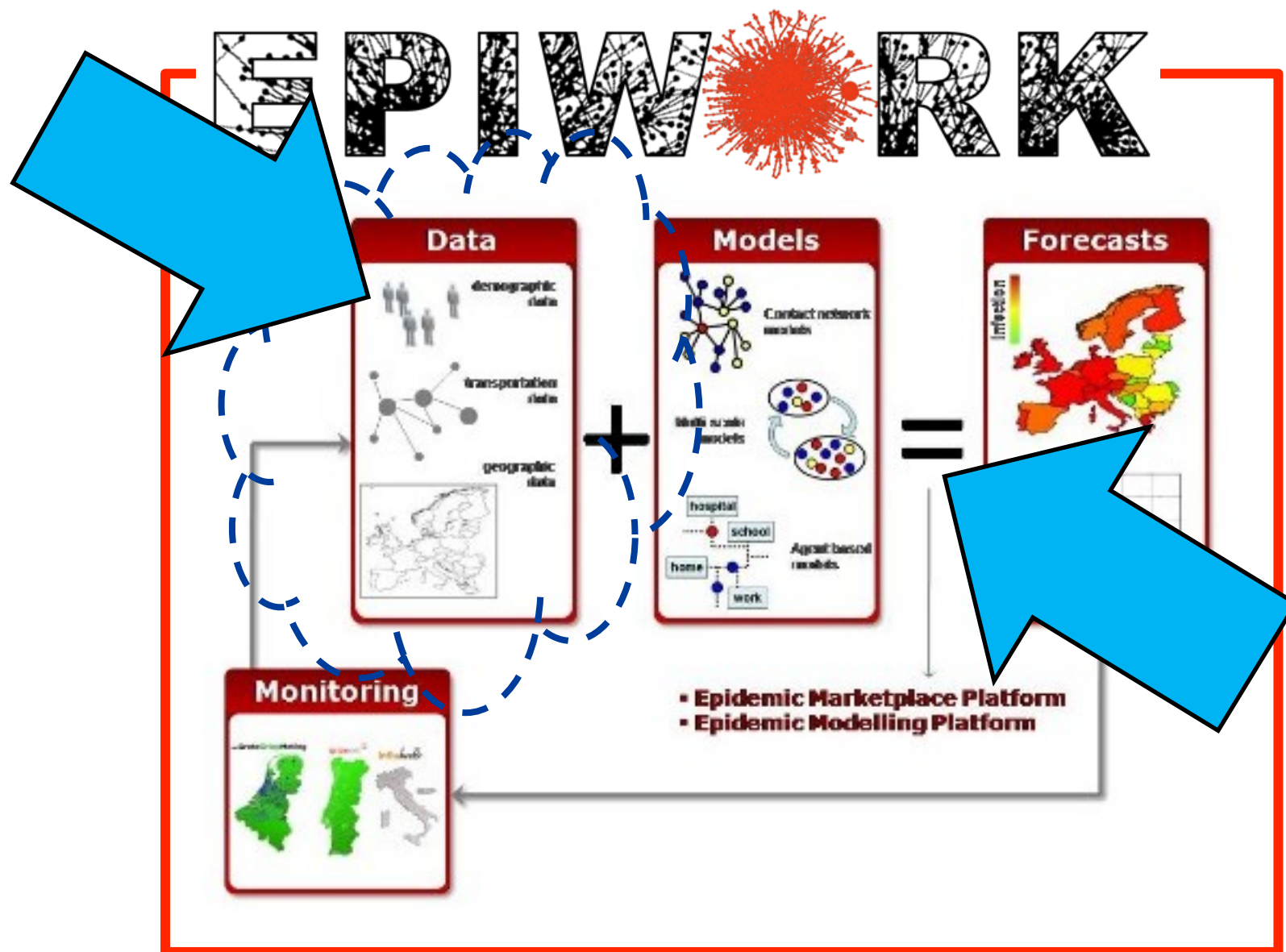
Torino 29th & 30th May, 2013

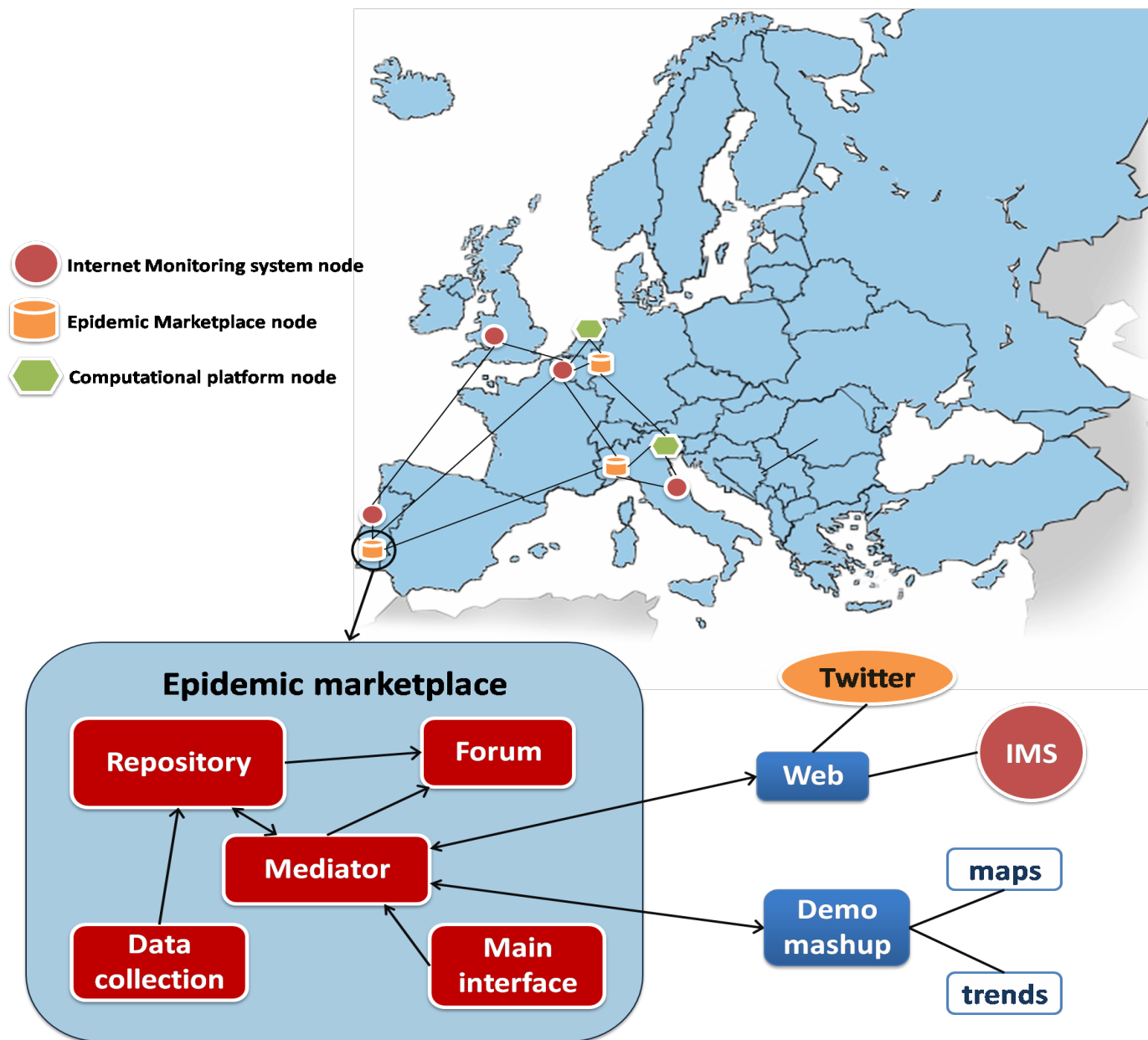
Mário J. Silva

FFCUL

mjs@inesc-id.pt



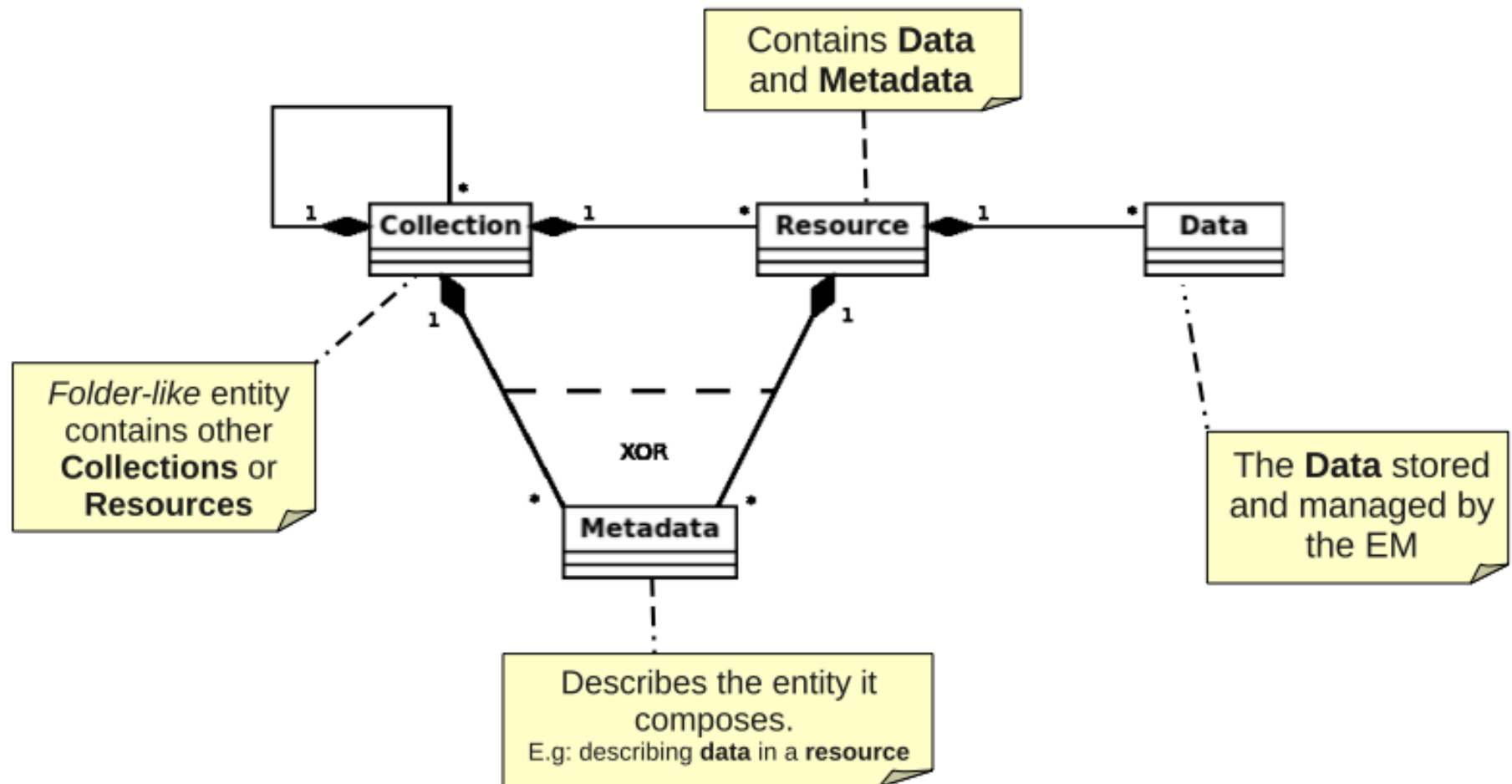




Epidemic Marketplace (EM)

1. **Catalogue of data sources** containing the metadata describing existing databases;
2. **Forum** to
 - publish information about data
 - seek modelers to collaborate with,
 - seek sources of data that could be of interest to their epidemiological modeling efforts;
3. **Mediating software** to automatically process queries to epidemiological data, harvest data, assemble datasets....

Conceptual Model of the EM Repository



Timeline of EM Releases

Date	Release	Event	Documents
Nov 2011	V1.0	(CO) Epiwork Meeting	D3.3
Jan 2012	V2.0	(PU) @ 1 st Epiwork Workshop	D3.4
May 2013	V3.0	@ 2 nd Epiwork Workshop	D3.6 (jul 2013)


[Home](#)

Resource Metadata

Meta-data	
Title	Twitter dataset H1N1 + Portugal 4-6-2009
Creator	LASIGE node of the Epidemic Marketplace
Subject	twitter message dataset
Description	<p>This dataset contains Twitter messages containing the words H1N1 and Portugal collected between 16-5-2009 and 3-6-2009, Information is a 7 columns relation, containing the following data:</p> <p>Column 1- keyword 1 (disease)- H1N1 Column 2- Keyword 2 (location)- Portugal Column 3- Source (Twitter) Column 4- Author of the message (user id) Column 5- The message body (evidence) Column 6- score Column 7- date (day and hour)</p>
Publisher	Epiwork – http://www.epiwork.eu
Format	text/tab-separated-values
Language	English, Portuguese
Contributor	Luis F Lopes, Joao M Zamite, Bruno C Tavares, Francisco M Couto, Fabricio Silva, Mario J Silva
Relation	Luis F. Lopes, João M. Zamite, Bruno C. Tavares, Francisco M. Couto, Fabrício Silva and Mário J Silva (2009). Automated Social Network Epidemic Data Collector. INForum informatics symposium.
Source	http://epiwork.di.fc.ul.pt/collector/
Coverage	Spatial: Portugal, Temporal: 2009-5-16 to 2009-6-3
Rights	Creative Commons Attribution-ShareAlike (CC BY-SA), http://creativecommons.org/licenses/by-sa/3.0/

**Initial Graphic
Style
Consortium use only**

Epidemic Marketplace

// Epidemic Marketplace

// Datasets

// Forum

// Web Services

// Data Collection

// Contact

// Learn More

Epidemic Marketplace

... a platform for integrating and sharing epidemiological data.

**V1.0 Graphic Style
Public in 2nd year**

Some features of the Epidemic Marketplace require user authentication. Please [register](#)

Datasets

Forum

Web Services

Data Collection

Announcements

2010.10.01 // Epimarketplace.net is now open to user registrations from the entire community. [\(See More\)](#)

[Browse](#)[Upload](#)[Request](#)

Latest Announcement

New upload

The new upload is now available at v2.epimarketplace.net/newupload

[See More](#)

Advanced Search

[Become a Curator](#)

A curator is a user that has some administrative permissions over the Epidemic Marketplace website. [See more](#)

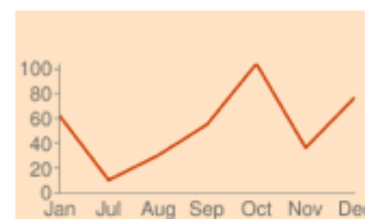
[Developers Corner](#)

You can access to our Web Services, Tools, Manuals and Significance. [See more](#)

100 Resources

[See More](#)

379 Visitors

[See More](#)[Recent Requests](#) (0)[Popular Requests](#) (0)[Recent Uploads](#) (0)

Website Tour

Most visited documents of the week

[Epi Info](#)

Luis Filipe Lopes

2010-09-17

Epi Info™ and a personal computer, epid ...

Subject: epidemic
Type: Software

[See more](#)

[cumEl Salvador](#)

Daniela Paolotti

2010-06-10

Subject: epidemic
Type: Dataset

[See more](#)

Announced at 3rd meeting
Try it at:
Epimarketplace.net



epidemic
marketplace

In 2009 there were 14,346 fatal cases of swine flu,
we stored the data

Examples: Epidemic, flu, h1n1

Search

Advanced ►

Latest resources



Mobility data

Seasonal transmission potential and activity peaks of the new influenza A(H1N1): Supporting Data

by Francisco Couto on 2013-05-15

Epidemiology



Event

Workshop on Digital Epidemiology

by Francisco Couto on 2013-05-15

Epidemiology



Epidemiological
data

HCoV-EMC infection in Calu-3 cells: Supporting data

by Francisco Couto on 2013-05-15

Epidemiology

Latest requests



Request for flu information

✓ Epidemiology

Announced HERE at 4th meeting
Try it at:
Epimarketplace.net

Request for additional data

✓ Epidemiology

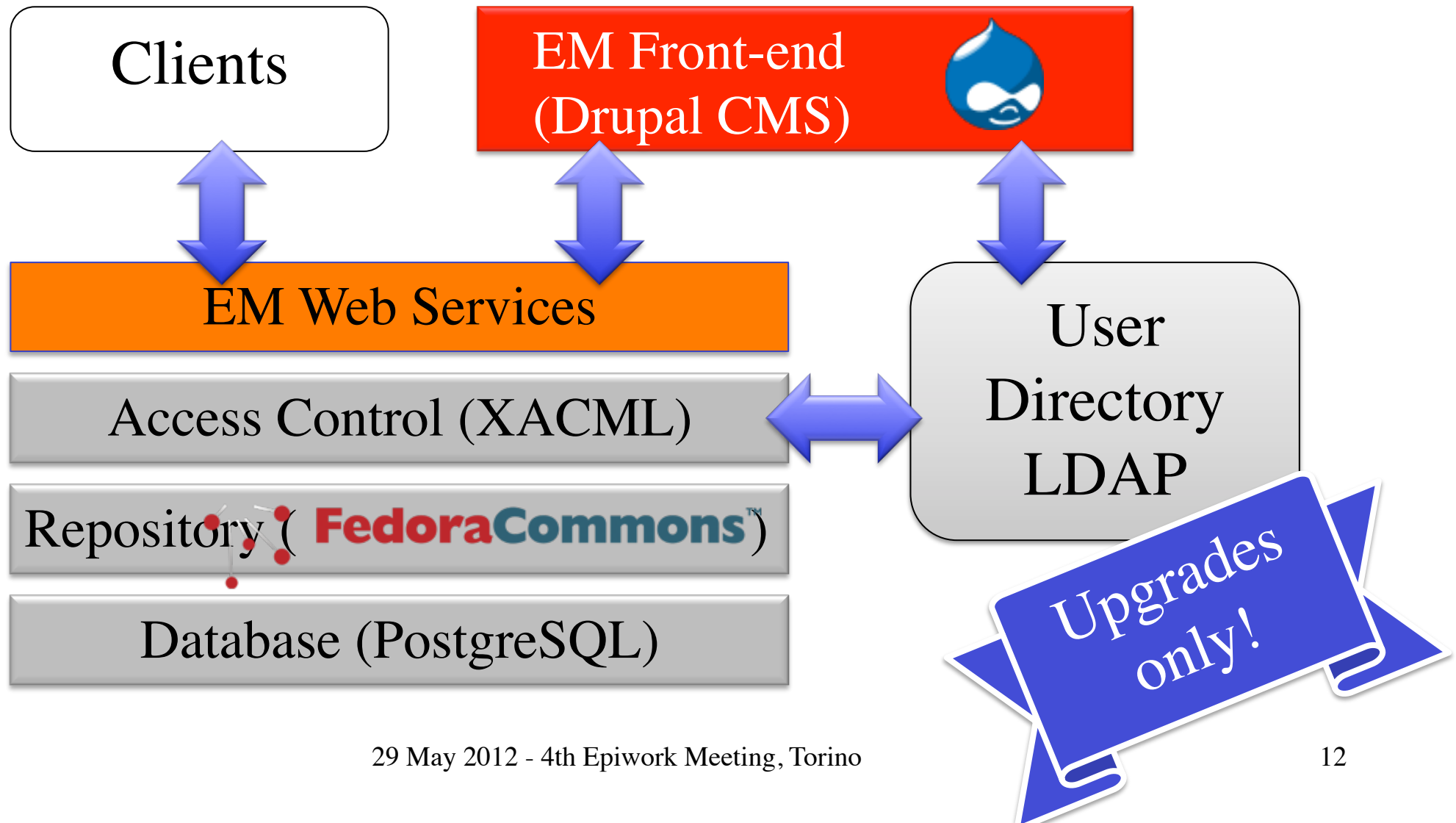


epidemic
marketplace

EM 3.0 Front-End

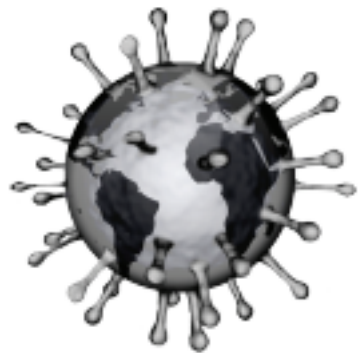
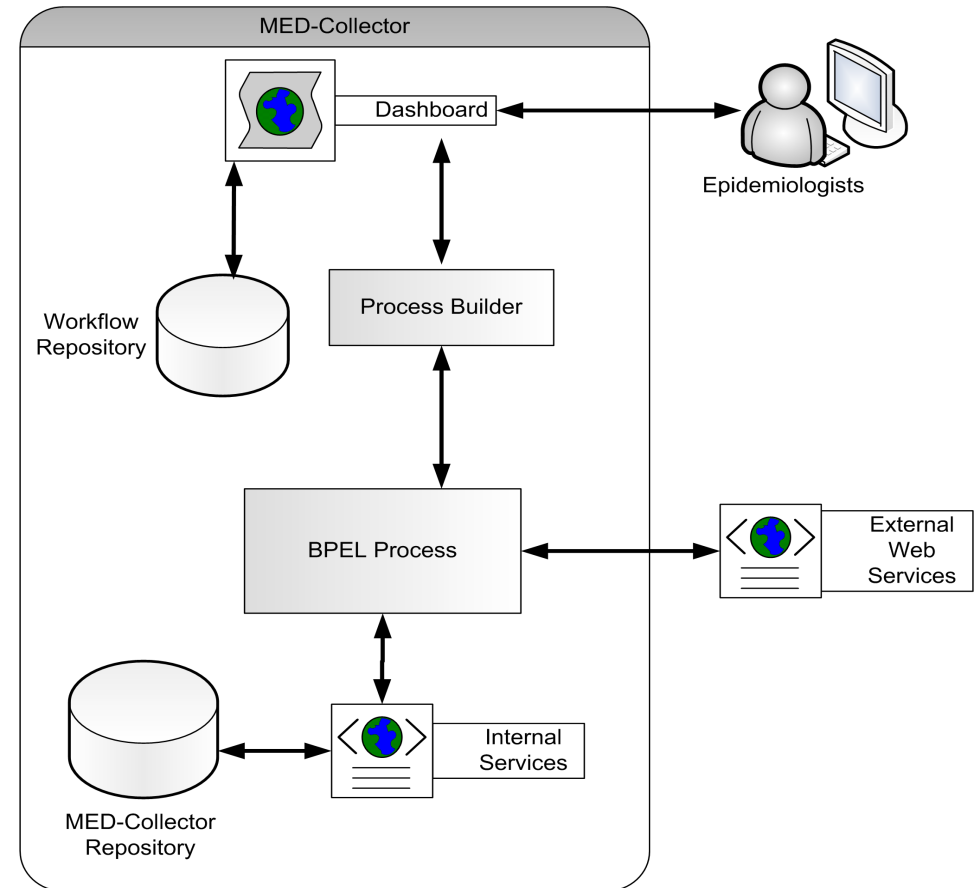
- Focus on usability and design
- Much Improved annotation interface
 - Support for manual/automatic annotation
- Curator portal: gone
 - The “collection” metaphor is enough

Software Components



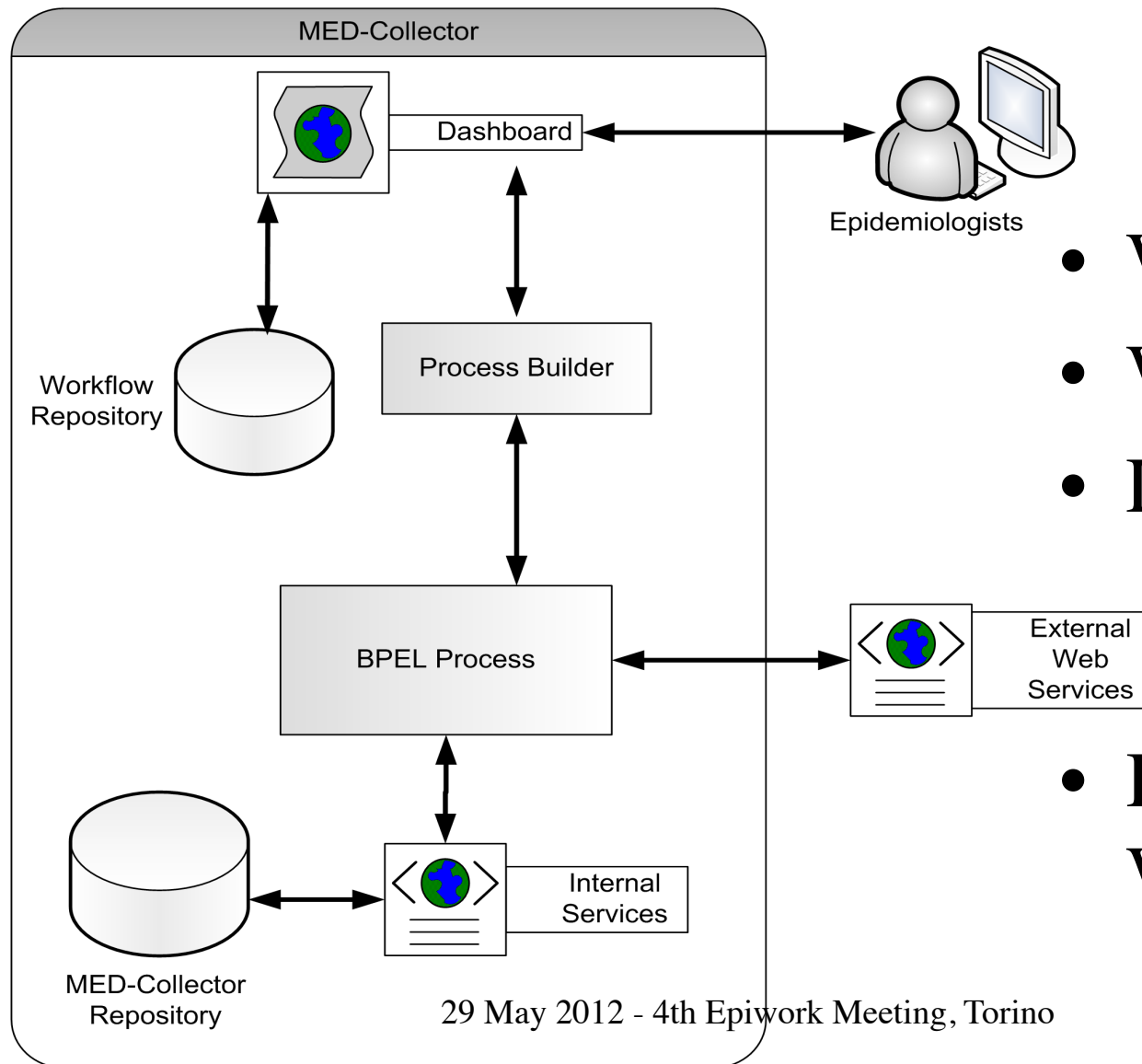
EM Clients

- MedCollector
- GleanViz



GLEaMviz Simulator
GLEaMviz.org

MEDCollector Architecture

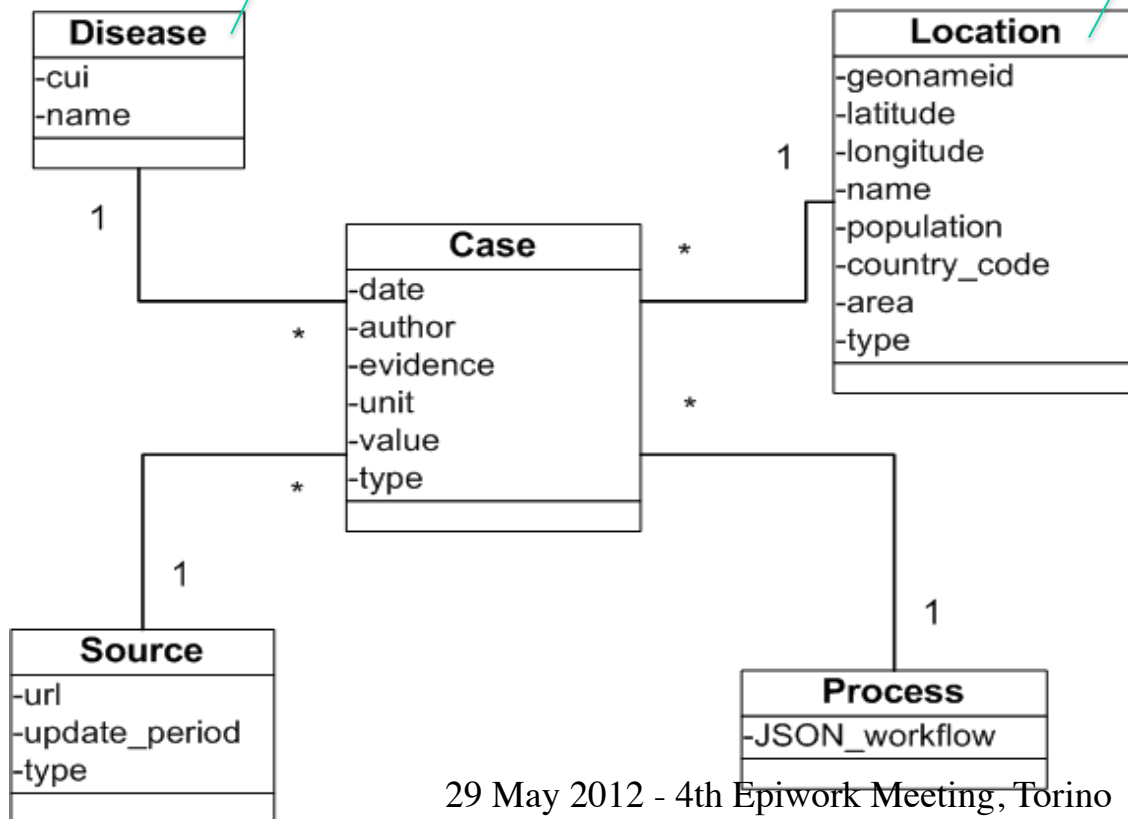


- Web Services
- Workflow Processes
- Local Storage
 - MEDCollector Repository
- Dashboard for Workflow Design

MEDCollector Data Model

UMLS: twitter searched
ion subset of 89 terms
related to “Disease or
Syndrome”

Geonames.org:
All Countries
and Capitals



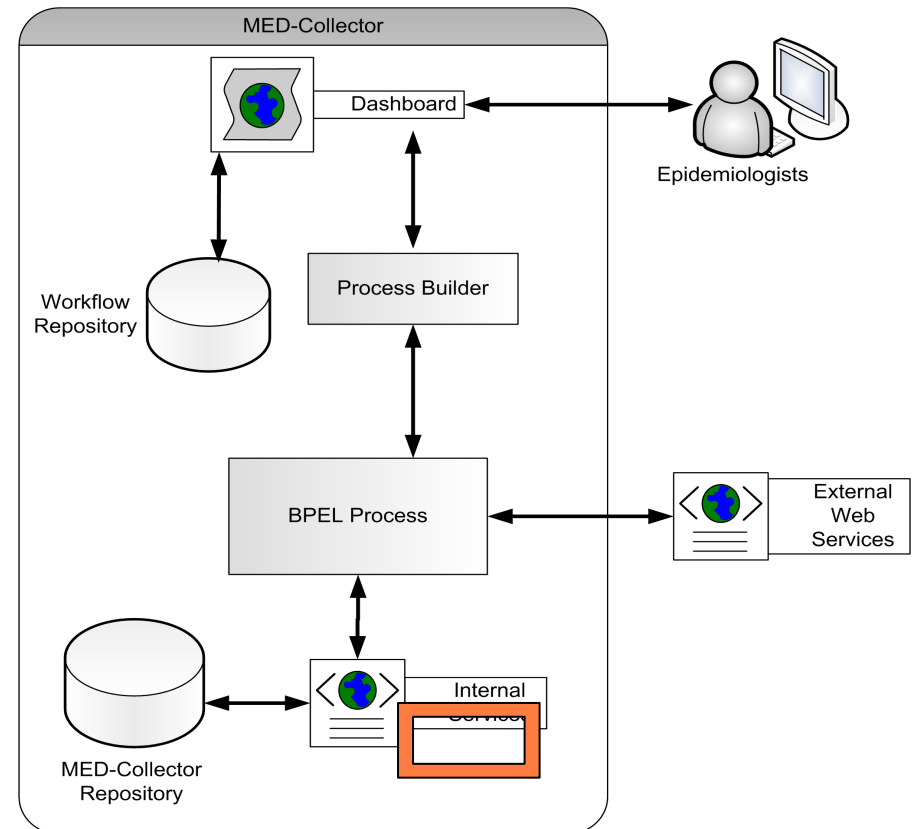
MEDCollector Services

❏ Data Collection Services

- ❏ The Query Selection Services
- ❏ Data Harvesting Services
- ❏ XML Transformation Services
- ❏ Database Loading Service

❏ Data Packaging Services

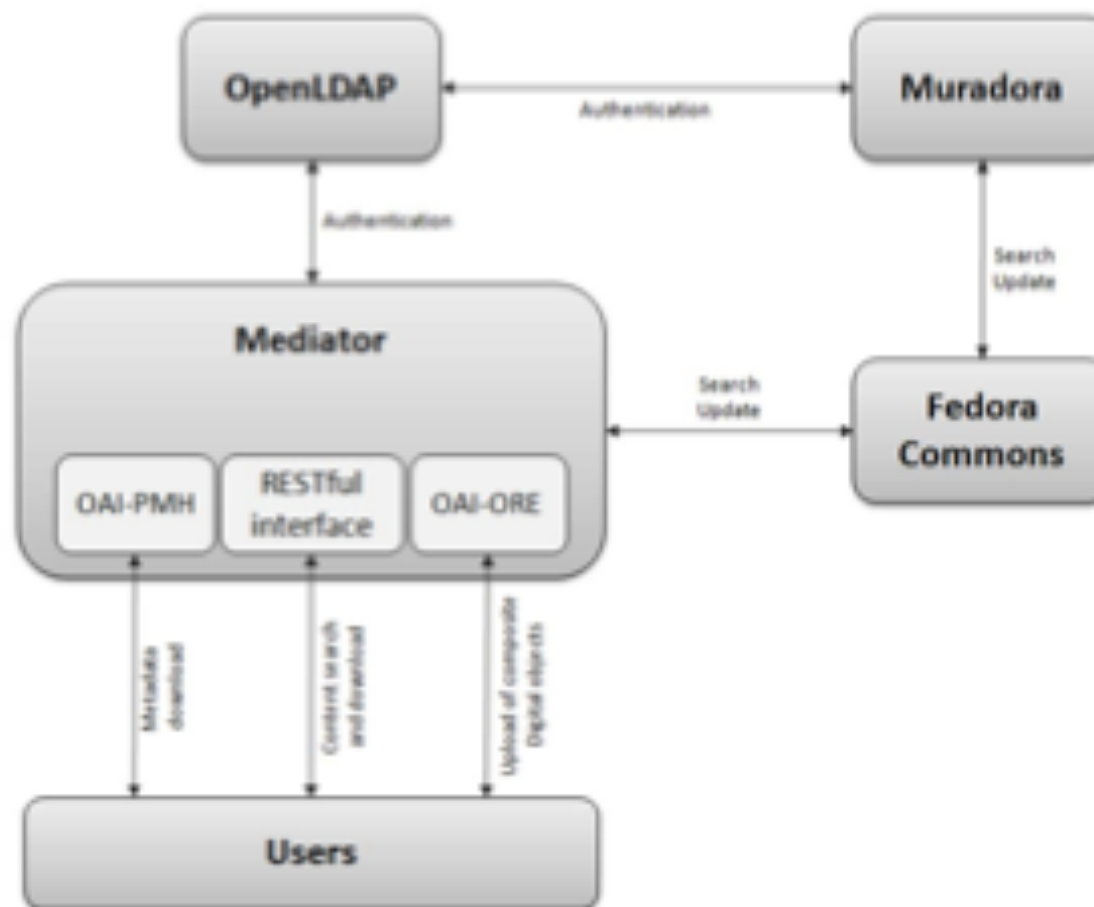
- ❏ To CSV



Mediator Web Services

7

- Upload
- Fetch
- Search



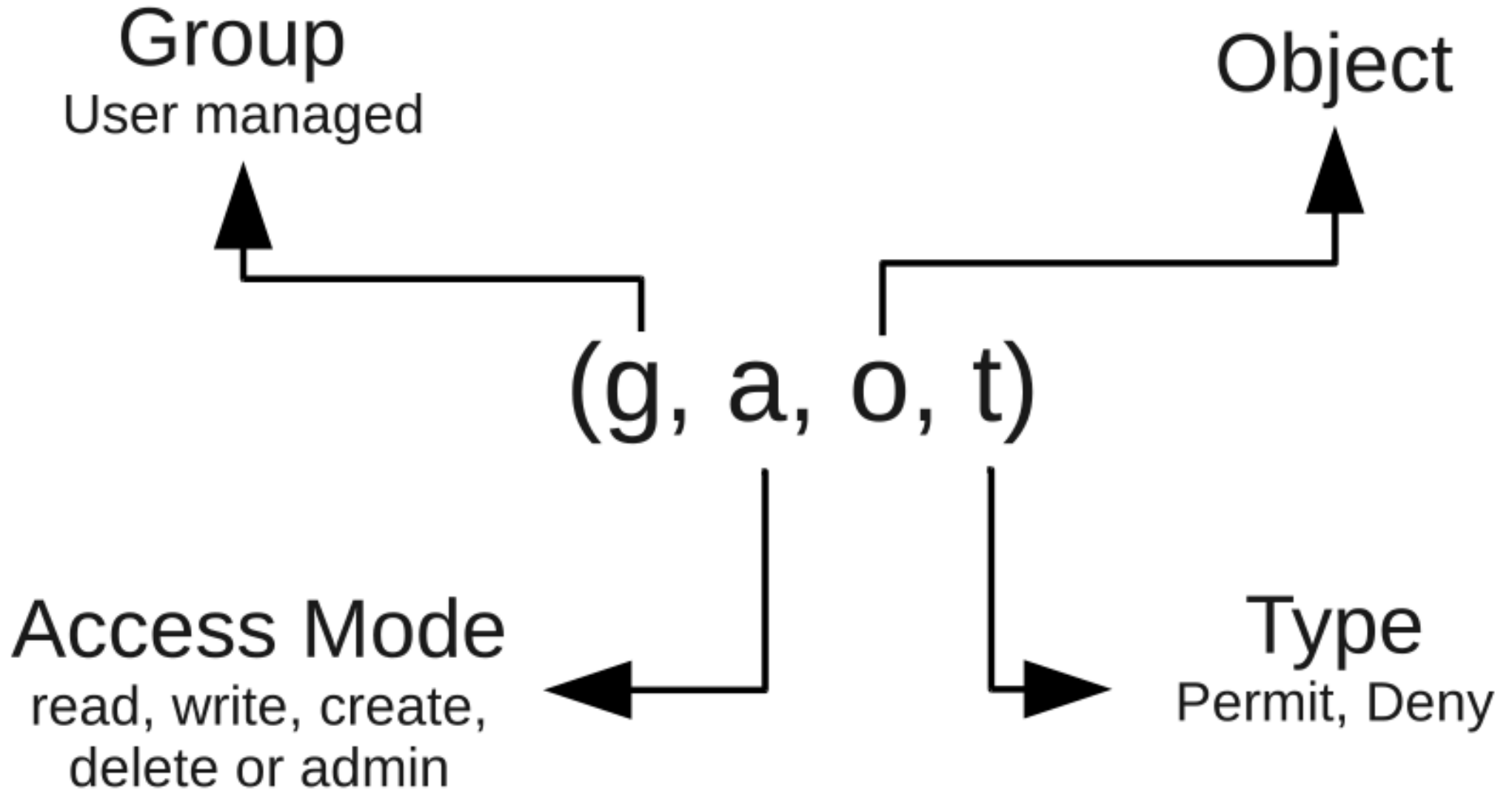
EM Web Services 2.0

- Same for EM front-end and remote clients
- URL of documentation: <http://>

Simple EM Client

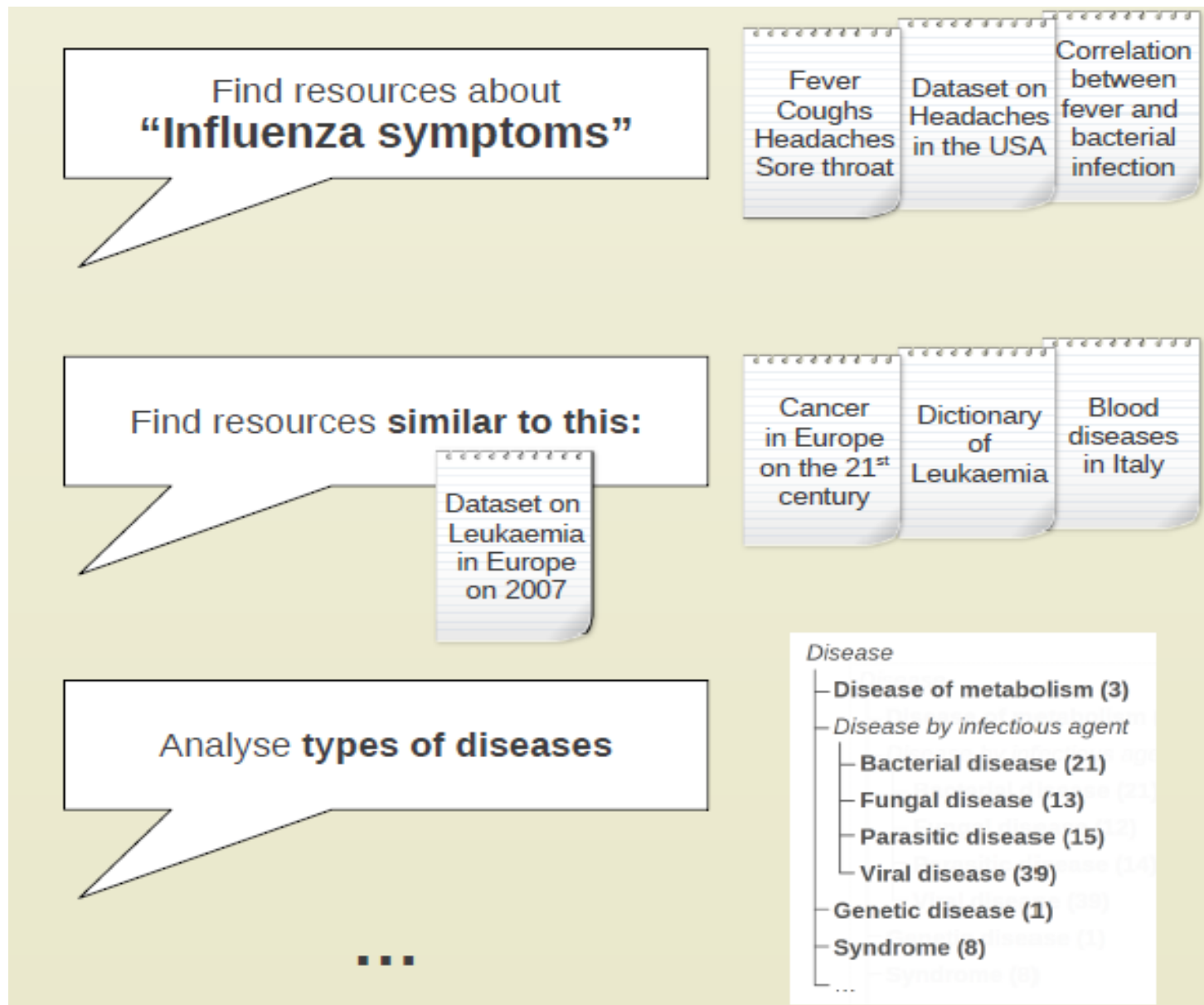
- Mapping of client filenames to EM resources (FC data streams)
- Operations: Check-out, check-in, synchronization
- Later-on: mapping of directories to collections and dataset synchronization

Access Control Model

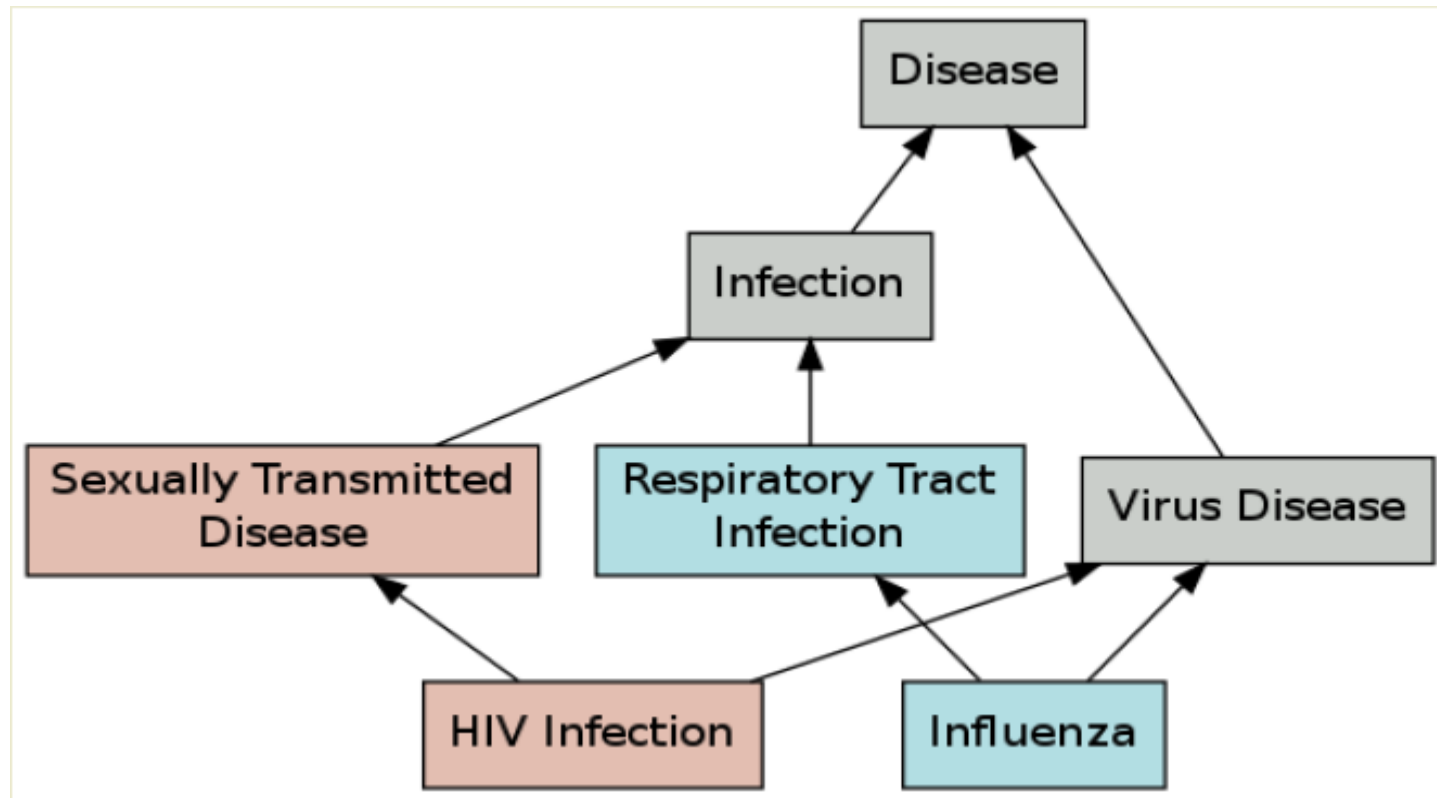


Access Control Model Implementation

- ??



NERO (see D3.5)



Epidemiology-
related ontologies

“Foreign” epidemic
resources

feedback

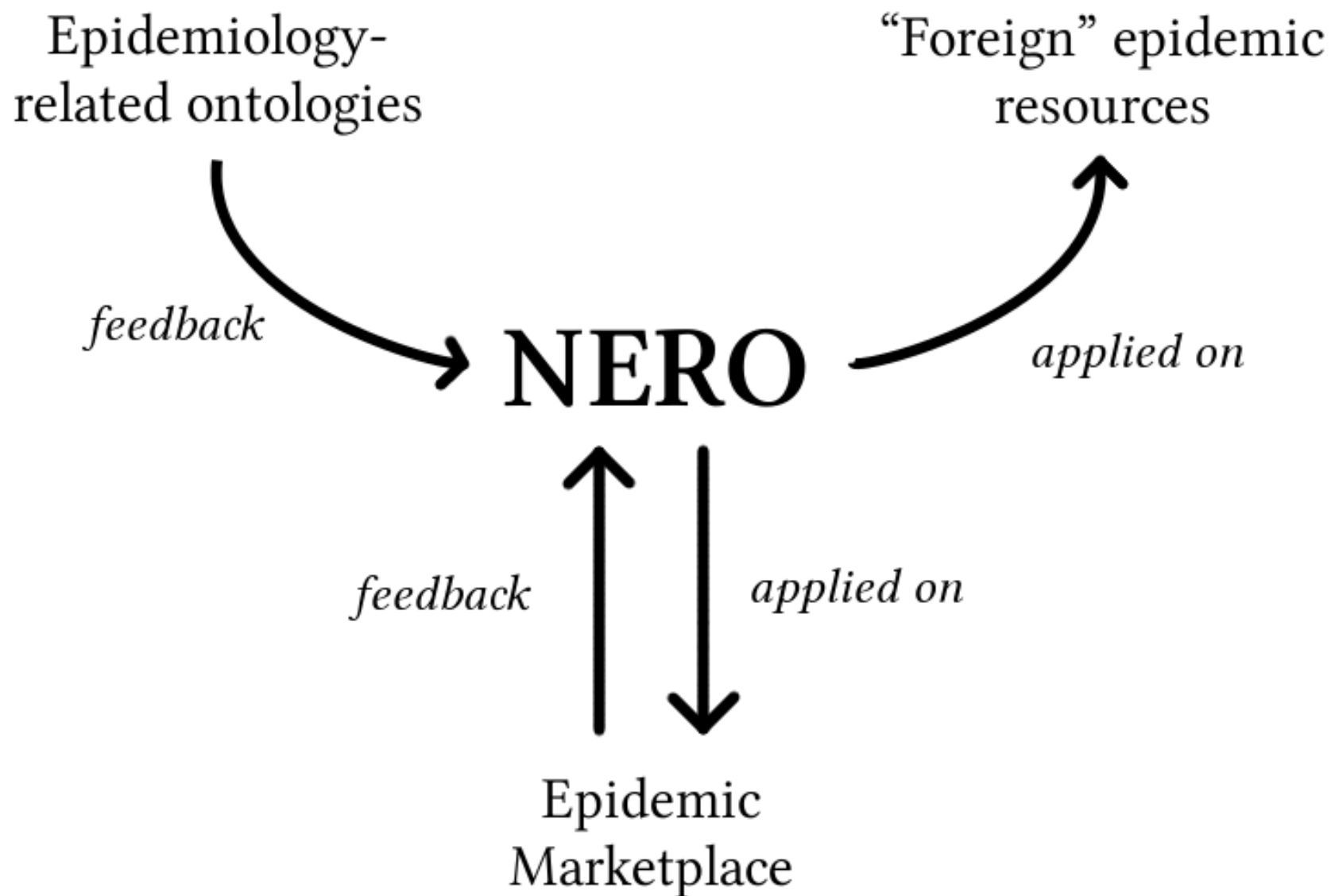
NERO

applied on

feedback

applied on

Epidemic
Marketplace

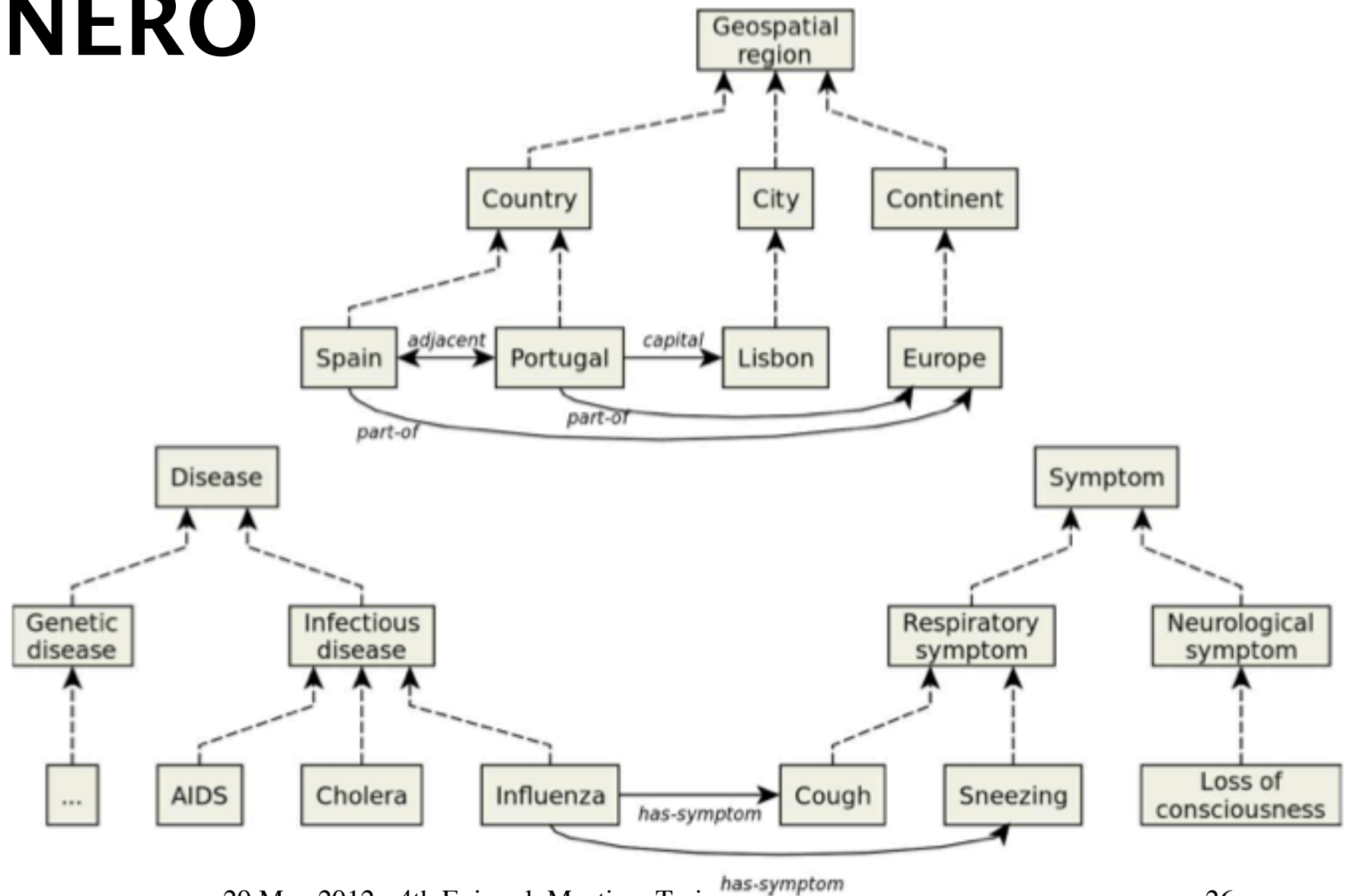




epidemic
marketplace

ONTOLOGY

NERO



Ontologies selected for NERO

- MeSH
- GeoPlanet
- ENVO
- TRANS
- CheBI

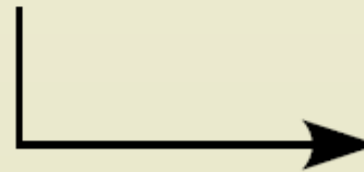
MeSH used to index articles in biomedical sciences (**general domain**)

GeoPlanet™ representation of the world **geography**

ENVO ontology for the **environment** domain

TRANS describes **transmission** of a pathogen from one host to another host

ChEBI ontology of **chemical compounds** focussed on biologically relevant entities



Upload resource	
Title:	Influenza in France
Description:	Contains data collected by an organisation
Disease:	Influenza
Host:	Human
Transmission:	Airborne
Environment:	Temperature, Humidity
Location:	France
Demography:	School years
...	

Resource:

Edit "Infections, medication use, and the prevalence of symptoms of asthma, rhinitis, and eczema in childhood.":

The Pre-load feature is based on [BioPortal Annotator](#) tool and NERO ontologies.

Note: you will have to submit changes after the pre-loading.

[Pre-load](#)

Biological Information

Diagnostic Method

Diagnostic test used to obtain data in the resource; or the protocol defined in the resource

Disease

Disease that is mentioned in the contents of the resource

Symptom

Symptom that is mentioned in the contents of the resource

Drug

Chemical compounds associated with in the resource, as cause and/or treatment of a disease

Other available actions

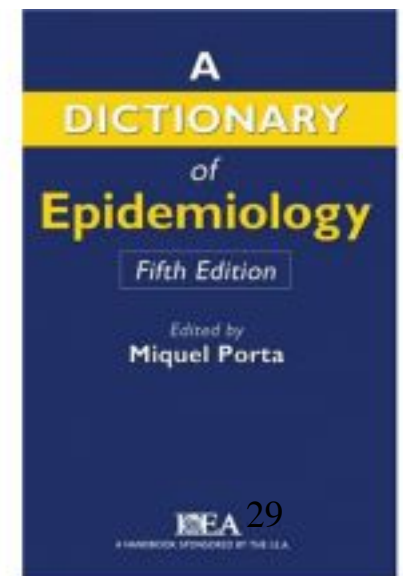
[Title and description](#)
[General information](#)
[Time and location](#)
[Epidemiological information](#)
[Authors](#)
[Source information](#)
[Bibliographic references](#)
[External identifiers](#)
[Resource's rights](#)
[Upload file](#)
[Edit Urls](#)
[VIEW RESOURCE ▶](#)

Set a relationship with
this resource

[← Pick resources](#)

Dictionary of Epidemiology

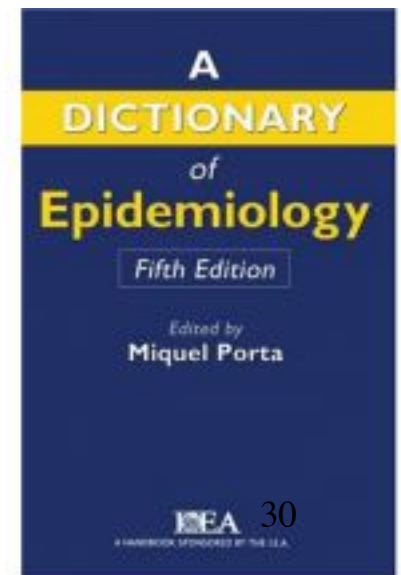
- Contacts with Miquel Porta (editor of the *Dictionary of Epidemiology*, 5th edition)
- Discussing the upgrade of the Dictionary of Epidemiology to an Epidemic Ontology
 - Will add semantic relationships among dictionary terms



Idea: Semantification of the Dictionary of Epidemiology

DISEASE MODEL - Quantitative simulation of the natural history of a disease (incidence, progression, prognosis, etc.) using epidemiological data. A public health model is population-based and used for planning and evaluating health services, whereas a clinical model is patient-based and used for patient care.

- Not so much a good idea after all



The Epidemiology Ontology (EO)

- **BASED** on the Dictionary of Epidemiology
- Focused on **Transmission**
 - OBO is welcoming
- JBMS paper, special issue on Ontology, in preparation

<https://code.google.com/p/epidemiology-ontology/>



epidemic
marketplace

EUROSURVEILLANCE STUDY

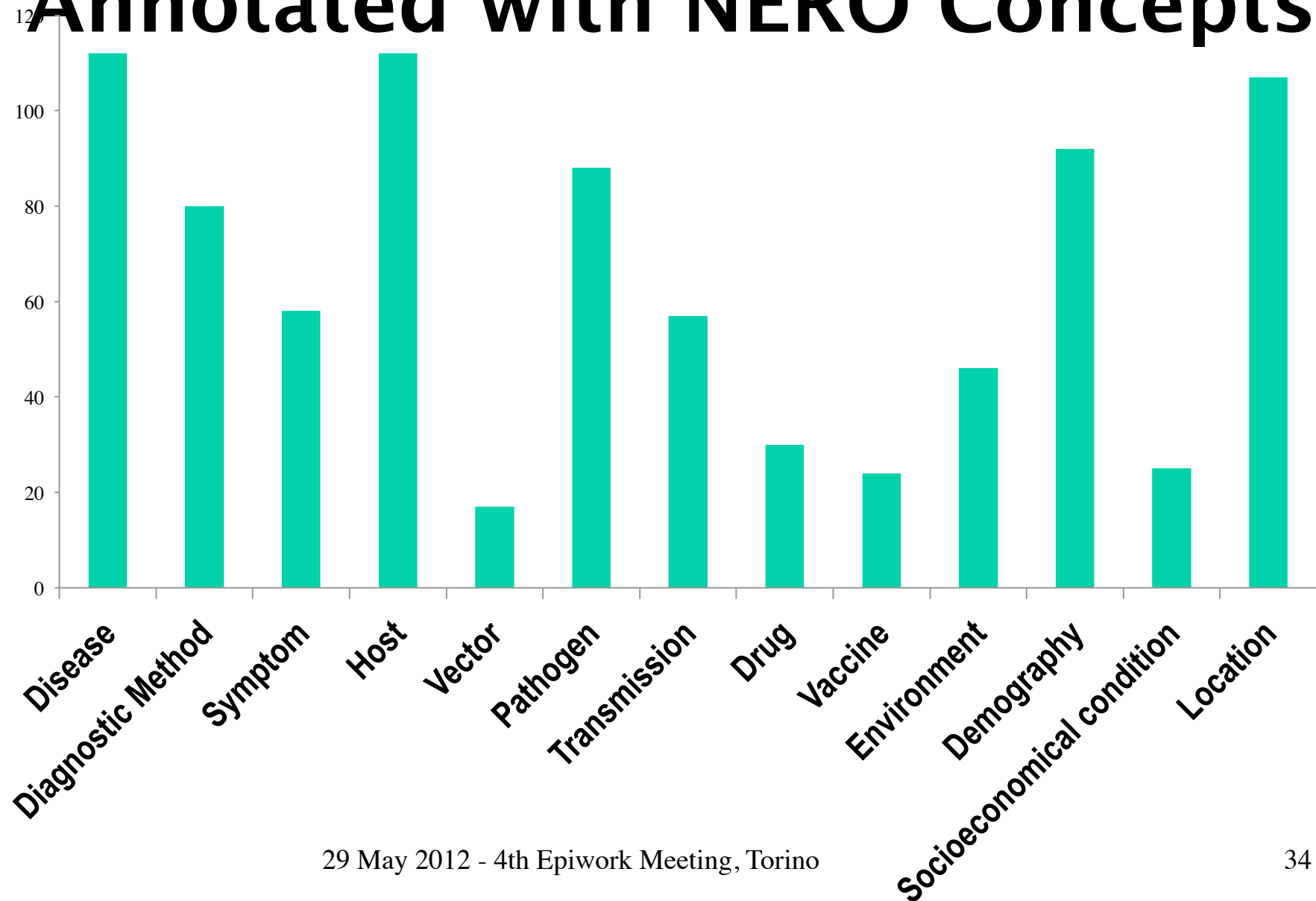


Eurosurveillance + JECH Study

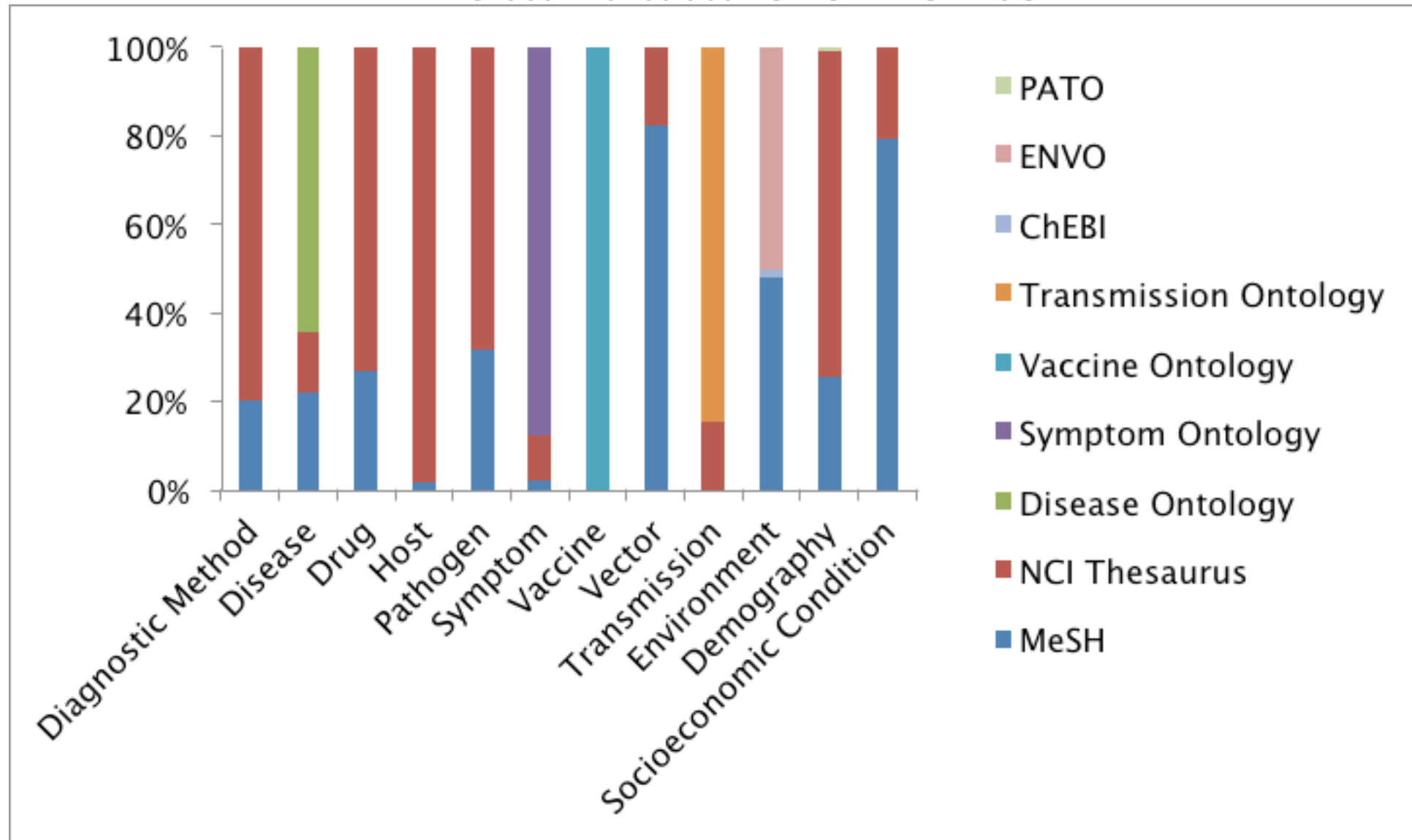
- Eurosurveillance: all (78) papers published between October 13, 2011 and August 16, 2012.
- JECH: 2005 to 2011, 34 papers.
- **112 publications.**
- All datasets and corresponding papers fully characterized in the Epidemic Marketplace

Papers

Annotated with NERO Concepts



Distribution of Ontology Terms in meta-data elements



Eurosurveillance Study: Conclusions

1. **None** of the 112 datasets **available from a repository**
2. Textual descriptions **provide the necessary information** for reasonably complete characterization of datasets with NERO concepts.
3. Having the datasets in public repositories, like the EM, would **stimulate production of epidemic data as linked open data** and enable meta-analyses in epidemiology
4. **Incentives** are needed to promote the production of structured metadata characterising epidemic datasets.



epidemic
marketplace

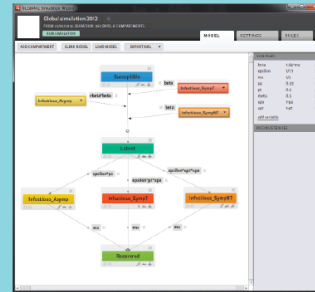
INTEGRATION



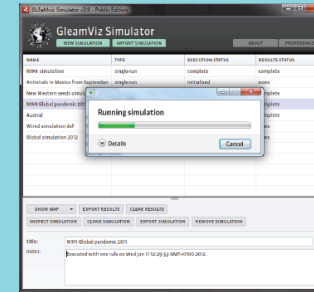
Visiting User



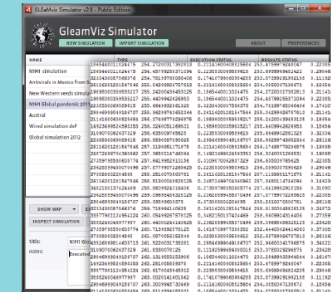
Epidemiologist



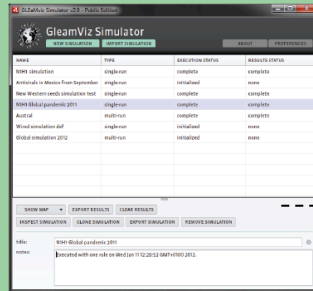
1. Design a new simulation



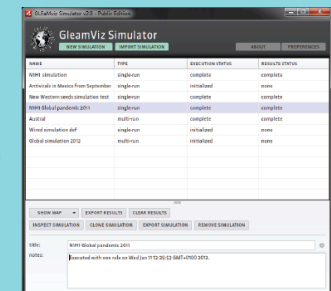
2. Run the simulation



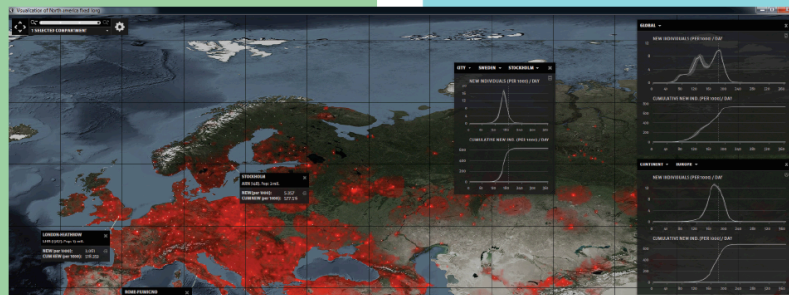
3. Get the output



1. Browse the epidemic marketplace



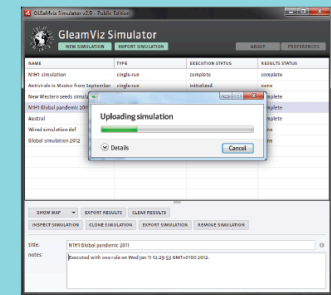
4. Browse the epidemic marketplace



2. Explore the results



6. Share the simulation



5. Upload the simulation

7. Explore the results

WP2 Integration



Search for... **Search** Advanced ►

Resource:

Synthetic contact matrices for Europe



uploaded by **Marco Ajelli** on 2013-05-30

Epidemiological data

0 0 4

Contact matrices by 1-year age brackets are available for 26 European countries, namely Austria, Bulgaria, Switzerland, Czech Republic, Germany, Denmark, Spain, Estonia, Finland, France, United Kingdom, Greece, Hungary, Ireland, Italy, Lithuania, Luxembourg, Latvia, The Netherlands, Norway, Portugal, Romania, Slovakia, Slovenia, Sweden, Cyprus.

File M.xls contains country-specific total matrices of adequate contacts by age.

File HSWR.xls contains setting-specific matrices of adequate contacts (household, school, workplace, general community).

In files M-symm.xls and HSWR-symm.xls there are the symmetric versions of total and setting-specific contact matrices, respectively.

PREVIEW

Unsupported format type for previewing

Uploader



Marco Ajelli

Organisation:
FBK

URL:
<http://www.fbk.eu>

Contact

Tags

Demography Epidemiology

License



To the extent possible under law, the person who associated CC0 with this work has waived all copyright and related or neighboring rights to this work.

WP5 Integration



Filter by

Disease

Influenza (11)

Country

Belgium (3)

Italy (3)

Netherlands (3)

Portugal (3)

Spain (1)

France (1)

United Kingdom (1)

Sweden (1)

UK (1)

Search

Search

Examples: Epidemic, flu, h1n1

Advanced ►

Order by: Title Uploader Date

11 results found



Epidemiological
data

Influenzanet Estimated Incidences for UK in 2012

by João Miguel Quintino de Moraes Zamite on 2013-01-30

Influenzanet is a system to monitor the activity of influenza-like-illness (ILI) with the aid of volunteers via the internet. These are the results of this survey for the country UK and the period 2012.

Epidemiology



Organisation

Influenzanet

by Luis Filipe Lopes on 2010-02-09

Epidemiology





epidemic
marketplace

DISSEMINATION

Events – 4th Period

- Participation in SemEval 2013
 - Recognition of drug names using ChEBI (one of NERO's ontologies)
 - Best (macro-averaged) F1 in the MEDLINE test dataset annotation



- ICBO 2012
 - 1 paper
 - Demoed the EM and NERO annotator



- Bioinformatics Open Days 2012. Braga, Portugal. Poster

29 May 2012 - 4th Epiwork Meeting, Torino

Publications – 4th Period

1. Tiago Grego, Francisco M. Couto 2013: Enhancement of Chemical Entity Identification in Text Using Semantic Similarity Validation. **PLOS ONE** 5(8), e62984.
2. João Ferreira, Daniela Paolotti, Francisco Couto, Mário J. Silva 2012: On the usefulness of ontologies in epidemiologic research and practice. **Journal of Epidemiology and Community Health**.
3. Catia Pesquita, Francisco M. Couto 2012: Predicting the Extension of Biomedical Ontologies. **PLOS Computational Biology** 9(8), e1002630.
4. João D. Ferreira, Francisco M. Couto, Semantic Similarity in the Biomedical Domain. Braga, Portugal. Poster @ Bioinformatics Open Days 2012. 2012
5. Francisco M. Couto, João D. Ferreira, João Zamite, Carlos Santos, Tiago Posse, Paulo Graça, Dulce Domingos, Mário J. Silva, The Epidemic Marketplace Platform: towards semantic characterization of epidemiological resources using biomedical ontologies. International Conference on Biomedical Ontologies (ICBO) 2012.
6. João D. Ferreira, Catia Pesquita, Francisco M. Couto, Mário J. Silva, Bringing epidemiology into the Semantic Web. International Conference on Biomedical Ontologies (ICBO) 2012.



epidemic
marketplace

TEAM & PLANS

FFCUL Team on WP3

- Mário J. Silva, Dulce Domingos, Francisco Couto, Cátia Pesquita
 - João Zamite, João Ferreira, Paulo Graça.
 - Tiago Posse, Carlos Sousa
- +
- Vera Carvalho, Hugo Santos (3 months)

WP3: status

- D3.6 due Jul 2013
 - Final report on the EM deployment

- Announcing the...



epidemic
marketplace

ToDo until JUL 2013...and beyond

- Short term
 - Complete D3.6
 - Contribute to PMR
 - Finish implementation of EM 3.0 user interface
 - Dissemination activities
 - 3+ submissions planned.
- Medium-long term



epidemic
marketplace

www.epimarketplace.net

In 2009 there were 14,346 fatal cases of swine flu,
we stored the data

Examples: Epidemic, flu, h1n1

Search

Advanced ►

Latest resources



Mobility data

Seasonal transmission potential and activity peaks of the new influenza A(H1N1): Supporting Data

by Francisco Couto on 2013-05-15

Epidemiology



Event

Workshop on Digital Epidemiology

by Francisco Couto on 2013-05-15

Epidemiology



Epidemiological
data

HCoV-EMC infection in Calu-3 cells: Supporting data

by Francisco Couto on 2013-05-15

Epidemiology

Latest requests



Request for flu information in Brasil

✓ Epidemiology



medcollector datasets

✓ Epidemiology



Data describing the population's infection network structure

✓ Epidemiology



Requesting additional data

✓ Epidemiology

29 May 2012 - 4th Epiwork Meeting, Torino

48

Todo list and planning (Torino, Dec 2010)

1. Populate Repository
2. New Front-end and updated components
(prepare market)
3. Access control policies & distributed authentication
4. Linked Epidemic Data
5. Ethics, Privacy and Anonimization
6. Replicate EM node

Todo list and planning (Pré Saint Didier, Jan 2012)

1. Populate Repository
2. New Front-end and updated components
(prepare market)
3. Access control policies & distributed authentication
4. Linked Epidemic Data
5. Ethics, Privacy and Anonimization
6. Replicate EM node

Scheduled Deliverables

		Year 1			Year 2			Year 3			Year 4		
		M4	M8	M12	M16	M20	M24	M28	M32	M36	M40	M44	M48
WP3	Information platform												
Task 1			D3.1										
Task 2			D3.1										
Task 3				D3.2		D3.3				D3.4			D3.6
Task 4										D3.4 D3.5			D3.6